Complete Sequencing of Mono–Deprotonated Peptide Nucleic Acids by Sustained Off–Resonance Irradiation Collision–Induced Dissociation

Jason W. Flora and David C. Muddiman*
Department of Chemistry, Virginia Commonweath University, Richmond, Virginia, USA

Complete peptide nucleic acids (PNAs) sequence information is obtained from the unimolecular decomposition of singly-charged PNA oligomers in the negative-ion mode using electrospray ionization coupled with Fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICR-MS) and sustained off-resonance irradiation collision induced dissociation. The 4-mers, *n*-CATT-c, *n*-AGCT-c, *n*-AACT-*c*, and *n*-acetylated-AACT-*c* and two 6-mers, *n*-AAAAAA-*c* and *n*-CCCCCC-c, were investigated to explore the unimolecular decomposition of mixed-nucleobase and homopolymer PNAs representing purine and pyrimidine oligomers, respectively. PNA decomposition is explored using a product-ion appearance curve and double resonance experiments. A decomposition mechanism for sequence ion formation (PNA amide bond cleavage) is proposed. (J Am Soc Mass Spectrom 2001, 12, 805–809) © 2001 American Society for Mass Spectrometry

Peptide nucleic acids (PNAs), first synthesized in 1991 by Egholm and coworkers [1], are nucleic acid analogs offering a wide variety of applications in both the medical and biochemical communities. Because of their potential widespread applications, a rapid and accurate method for the characterization of PNA structures as well as thermal degradation and biotransformation products (i.e., in vitro and in vivo metabolism) is demanded. However, to date, gas-phase sequencing of these important biomolecules has been met with little success [2].

An experimental and computational investigation of singly-charged protonated PNAs, *n*-AACT-*c* and *n*-acetylated-AACT-*c* (the *n*-terminus and the *c*-terminus are analogous to the oligonucleotide 5′ and 3′ ends, respectively), using electrospray ionization [3] coupled with Fourier transform ion cyclotron resonance mass spectrometry [4] (ESI-FTICR-MS) illustrated that dissociation of these mono-protonated oligomers does not provide sufficient sequence information due to proton sequestering at the highly basic *n*-terminus [2c]. Three stages of MS using sustained off-resonance irradiation collision-induced dissociation [5] (SORI-CID) produced the non-sequence ions (described by a previously proposed nomenclature) [6] M–H_2O, M–$2H_2O$, M–H_2O–B_n(A), M–H_2O–$B_3(C)_{II}$, M–H_2O–$B_3(C)_{II}$–B_n(A) and the sequence ions w_3' and w_3'–H_2O [2c].

While extensively exploring the preparation of polymerase chain reaction (PCR) products for ESI-FTICR-MS in our laboratory [7], we have attempted the incorporation of PNAs to prevent re-annealing of single-stranded oligonucleotides [8]. This research unexpectedly revealed that PNAs provide an intense signal in the negative-ion mode under the appropriate ESI solution conditions. This is counter-intuitive because the PNA structure does not suggest that it would efficiently form negative-ions. The observation of successful deprotonation of PNAs and the interest in gas-phase sequencing of these oligonucleotide analogues demanded further investigation.

The 4-mers, *n*-CATT-c, *n*-AGCT-c, *n*-AACT-*c*, and *n*-acetylated-AACT-*c* and two 6-mers, *n*-AAAAAA-*c* and *n*-CCCCCC-c, were investigated to explore the unimolecular decomposition of mixed-nucleobase and homopolymer PNAs representing purine and pyrimidine oligomers, respectively. The PNAs were electrosprayed from a solution of 60:20:20 acetonitrile:isopropanol:10 mM ammonium acetate with piperidine and imidazole at final concentrations of 20 mM [9]. ESI-FTICR-MS conditions have been discussed elsewhere [7].

Figure 1 shows the product-ion spectra for the mono-deprotonated mixed-base oligomers produced by SORI-CID carried out in a Penning ion trap. All SORI-CID experiments were performed for 0.25 s, at 1000 Hz below the cyclotron frequency of the precursor ion, at

Published online May 10, 2001

Address reprint requests to Dr. D. C. Muddiman, Department of Chemistry, Virginia Commonwealth University, 1001 West Main Street P.O. Box 842006, Richmond, VA 23284. E–mail: dcmuddim@saturn.vcu.edu

*Affiliate Appointment in Biochemistry and Molecular Biophysics, and member, Massey Cancer Center, also in Virginia Commonwealth University.

1044-0305/01/$20.00
PII S1044–0305(01)00259–8

Received February 26, 2001
Accepted March 5, 2001

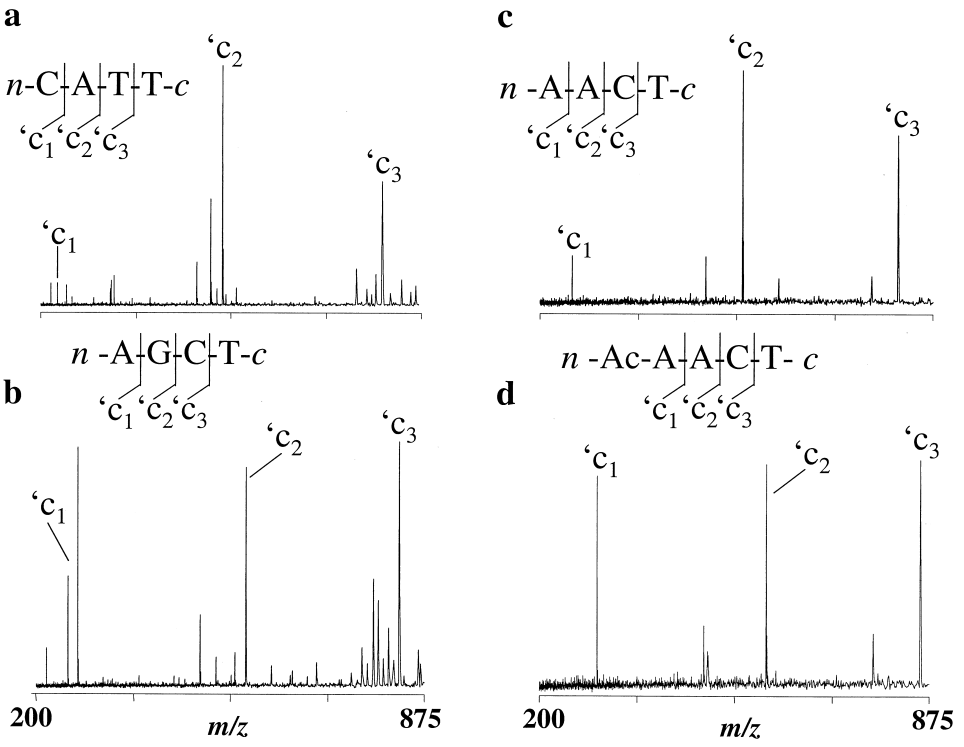


Figure 1. Product-ion spectra produced by SORI-CID for all mixed base PNA 4-mers resulting in *n*-terminal sequence ions, $'c_n$ ions. (**a**) *n*-CATT-*c* with SORI-CID carried out with an average kinetic energy in the laboratory frame of reference (E_{lab}) of 9.5 eV. (**b**) *n*-AGCT-*c* with E_{lab} of 16.5 eV. (**c**) *n*-AACT-*c* with E_{lab} of 11.6 eV. (**d**) *n*-acetylated-AACT-*c* with E_{lab} of 16.1 eV.

maximum pressure of $\sim 10^{-5}$ torr ($\sim 3.2 \times 10^{11}$ molecules/cm^3), with argon as the collision gas. All SORI-CID E_{lab} calculations use the cylindrical geometry factor ($\beta_{cylinder}$) of 0.86738, magnetic field strength of 4.655 tesla, and distance between excite plates of 6.05 cm.

Dissociation of these singly-charged PNAs, Figure 1, results in only sequence ions that contain the *n*-terminus, $'c_n$ ions, corresponding to cleavage of each peptide bond between the PNA residues (the subscript "n" signifies which peptide bond was cleaved). Figure 1a shows the sequence ions produced from SORI-CID of *n*-CATT-*c*. Figure 1b shows the sequence ions produced for the 4-mer PNA, *n*-AGCT-*c*, containing one of each nucleobase: adenine, guanine, cytosine, and thymine. Figures 1c and 1d show sequence information that is obtained for the two previously investigated PNAs, *n*-AACT-*c* and *n*-acetylated-AACT-*c*[2c].

Figure 2 shows the product-ion spectrum of the adenine homopolymer, *n*-AAAAAA-*c*. Six product-ions are observed, ions resulting from the cleavage of each peptide bond as well as an ion containing neutral base loss and the loss of water. Complete sequence information was also accomplished for the homo-pyrimidine, *n*-CCCCCC-*c*, (data not shown).

Figure 2 also demonstrates the internal calibration of PNA product-ions using polyethylene glycol with an average molecular weight of 1000 Da (PEG_{1k}) as the internal calibrant. Internal calibration of product-ion spectra was accomplished using a dual electrospray source [10] and hexapole accumulation [11]. External calibration of this data results in as much as one to two orders of magnitude greater mean mass error (ppm) than internally calibrated data, depending on the calibration file, which can often complicate product-ion identification. Clearly, internal calibration was not required for this system, however, it can be critical in more complex product-ion spectra (e.g., larger PNAs which are currently under investigation).

Figure 3 shows relative ion abundance as a function of SORI-CID E_{lab} ranging from 0 to 51.4 eV for *n*-AAAAAA-*c*. As the average kinetic energy increases (>3 eV) the relative abundance of M-H_2O ions reaches a maximum and all the *n*-terminal sequence ions, $'c_n$ ions (n = 1-5), begin to form as well as the nonsequence ion corresponding to neutral base loss, M-H_2O- B_n(A).

Determination of the mechanism by which product-ions are formed, however, is speculative at best when using fragmentation efficiency plots alone. Double-resonance experiments have been utilized to afford insight into the origin of each product-ion that is produced from tandem mass spectrometry measurements (MS^n). These experiments involve continuous ejection of a specific product-ion produced during SORI-CID by the application of an on-resonance superposition waveform. The persistence of the smaller product-ions indicates that their formation does not involve the ejected ion.

Double-resonance experiments were conducted on

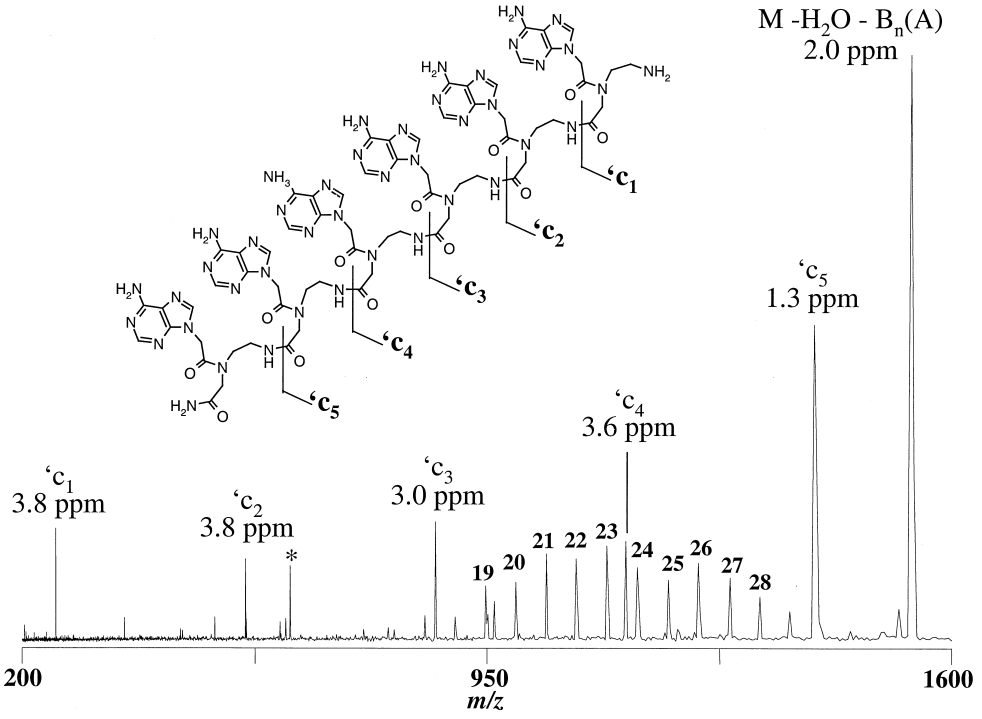


Figure 2. SORI-CID (single acquisition) of the adenine homopolymer 6-mer PNA at E_{lab} of 3.7 eV. The polyethylene glycol peaks used for internal calibration (average molecular weight of 1000 Da) are labeled according to oligomer numbers. Mass errors after internal calibration are listed in ppm and mean error is 2.9 ± 1.6 ppm (± confidence internal of the mean at a 99% confidence level). * = Noise peak.

the adenine homo-polymer to elucidate mechanistic pathways. The decomposition pathway $M \rightarrow MH_2O \rightarrow {}'c_n$ was explored by simultaneous ejection of $M\text{-}H_2O$ ions with a superposition waveform on-resonance with the ion's cyclotron frequency during SORI-CID. The ${}'c_n$ ions were still observed at low intensity, indicating that water loss is not required for the formation of sequence ions. Decomposition pathways such as $M \rightarrow {}'c_5 \rightarrow {}'c_4$ were also investigated by separate double-resonance experiments ejecting each product-ion.

Figure 4a shows the single-acquisition product-ion spectra of the adenine homo-polymer at an E_{lab} of 7.5 eV. All sequence ions are present, ${}'c_n$ as well as $M\text{-}H_2O$ and $M\text{-}H_2O\text{-}B_n(A)$. Figure 4b shows the product-ion spectrum resulting from SORI-CID of *n*-AAAAAA-*c* with concurrent ejection of the ${}'c_5$ product-ion with an on-resonance superposition waveform. The *m/z* of the precursor ion, SORI-CID excite frequency, and ${}'c_5$ ejection frequency (*m/z*) are all indicated. The presence of ${}'c_{1\text{-}4}$ product-ions indicates that the formation of these sequence ions are not dependent on the formation of ${}'c_5$. To determine if additional dissociation pathways are derived directly from the precursor ion or through an intermediate, additional double resonance experiments were performed. Specifically, off-resonance excitation of the precursor ion was conducted while sequentially ejecting the ${}'c_2$, ${}'c_3$, ${}'c_4$, and $M\text{-}H_2O\text{-}B_n(A)$ ions. In each individual experiment, only the purposefully ejected ions were absent from the product-ion spectra, therefore, establishing a direct pathway between the molecular ion and each individual product-ion (i.e., formation of intermediates) is not required.

Figure 3 illustrates that the initial water loss of PNAs clearly has the lowest energy of activation (excluding some of the ${}'c_1$ population) of all the PNA product-ions. This dehydration mechanism has been previously proposed and involves the charge remote dehydration of the primary amide at the *c*-terminus resulting in the formation of a nitrile [2c, 12]. Thermal decomposition of primary amides resulting in the formation of nitriles have been previously demonstrated using infrared spectroscopy, gas chromatography Fourier transform infrared spectroscopy, and gas chromatography mass spectrometry [12]. All PNA sequence ions contain the *n*-terminus and no site of dehydration, thus supporting the fact that the initial water loss occurs only at the *c*-terminal amide.

A mechanism for the formation of the *n*-terminal sequence ions, Scheme 1, involves deprotonation of the α-carbon of the carbonyl participating in the peptide bonds. The α-hydrogens are weakly acidic due to the electron withdrawing nature of the carbonyl oxygen. The α-carbon is less electronegative than the α-nitrogen, suggesting that this induced acidity is greater for the carbon. Subsequently, the α-carbon is more susceptible to proton loss. Upon deprotonation, the carbonyl group stabilizes the negative charge through enolate formation. Double-bond formation, proton transfer to the

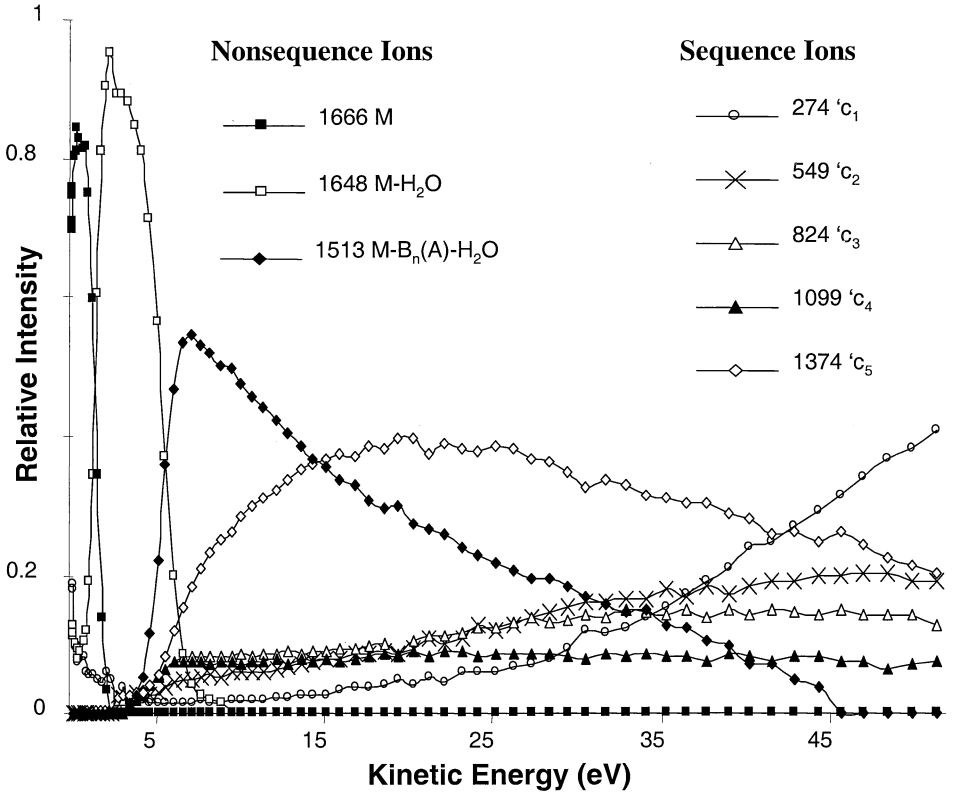


Figure 3. Average kinetic energy in the laboratory frame of reference verses product-ion relative intensity for the 6-mer adenine homopolymer. Masses and product-ion designations are labeled according to plot symbols.

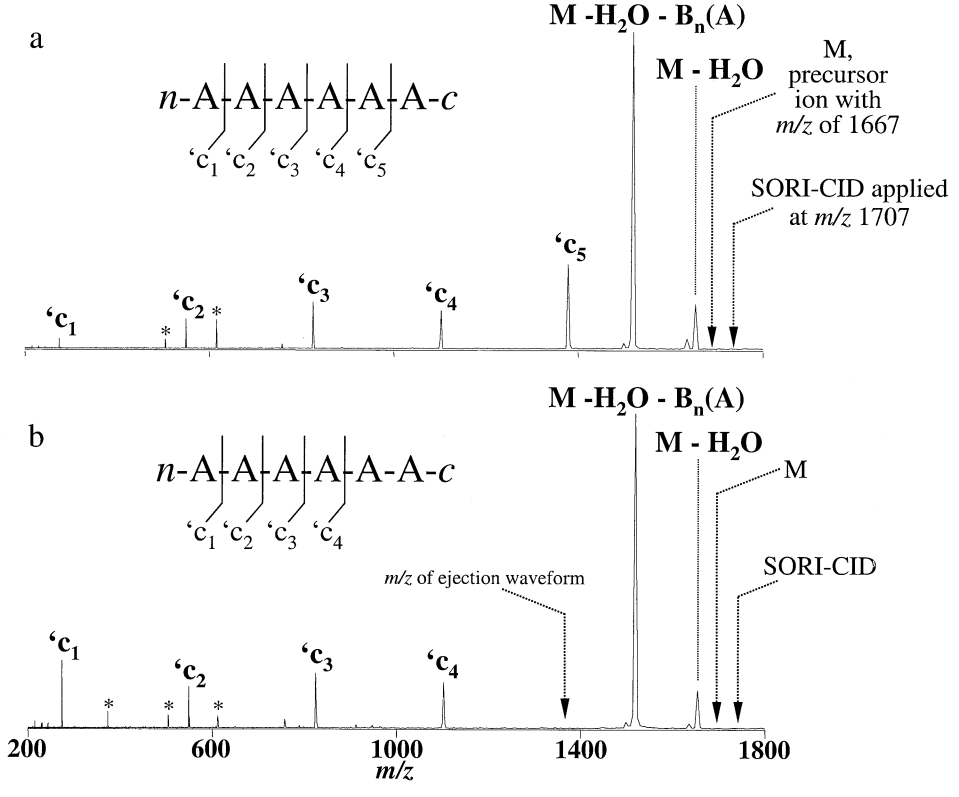
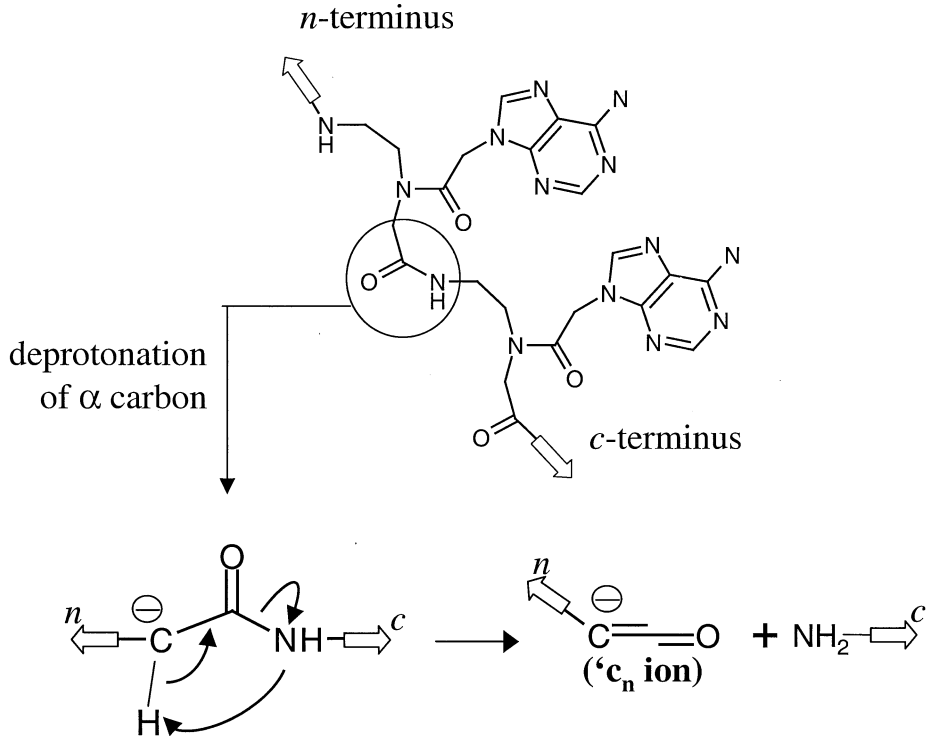


Figure 4. (**a**) SORI-CID (single acquisition) 1000 Hz off-resonance, corresponding to m/z = 1707, from the precursor ion M at m/z = 1667. (**b**) Double resonance experiment ejecting the sequence ion $'c_5$.



Scheme 1

basic nitrogen, and peptide bond cleavage transpires resulting in a mono-deprotonated ketene, $'c_n$ ion, and a neutral primary amine (*c*-terminal fragment).

To conclude, negative-ion dissociation of PNAs can be an effective strategy for the gas-phase sequencing of singly-charged PNAs. Complete sequence information was achieved for all mono-deprotonated PNAs investigated. Deprotonated PNAs have greater charge delocalization than protontaed PNAs[2c] which permits the formation of all *n*-terminal sequence ions. All dissociation channels are independent and the peptide bond cleavage mechanism has been proposed.

Acknowledgments

The authors express their appreciation for the financial support generously provided by the National Institute of Health and the Department of Chemistry, Virginia Commonwealth University.

References

1. (a) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991,** *254*, 1497–1500. (b) Egholm, M.; Buchardt, O.; Nielsen, P. E.; Berg, R. H. *J. Am. Chem. Soc.* **1992,** *114*, 1895–1897. (c) Egholm, M.; Nielsen, P. E.; Buchardt, O.; Berg, R. H. *J. Am. Chem. Soc.* **1992** *114*, 9677–9678. (d) Egholm, M.; Behrens, C.; Christensen, L.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Chem. Soc. Chem. Commun.* **1993,** 800–801.
2. (a) Takao, T.; Fukuda, H.; Coull, J.; Shimonishi, Y. *Rapid Commun. Mass Spectrom.* **1994,** *8*, 925–8. (b) Shampine, L. J.; Ding, J.; Davis, R. G.; Anderegg, R. J. *ASMS Conference on Mass Spectrometry and Allied Topics*; Chicago, IL, June 1994; pp 114–115. (c) Flora, J. W.; Shillady, D. D.; Muddiman, D. C. *J. Am. Soc. Mass Spectrom.* **2000,** *11*, 615–625.
3. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989,** *246*, 64–71.
4. Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974,** *25*, 282–283.
5. Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991,** *246*, 211–225.
6. Flora, J. W.; Muddiman, D. C. *Rapid Commun. Mass Spectrom.* **1998,** *12*, 759–762.
7. (a) Hannis, J. C.; Muddiman, D. C. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 323–330. (b) Muddiman, D. C.; Null, A. P.; Hannis, J. C. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 1201–1204. (c) Null, A. P.; Hannis, J. C.; Muddiman, D. C. *The Analyst (Special Issue on Biological Mass Spectrometry)* **2000,** *125*, 619–625. (d) Hannis, J. C.; Muddiman, D. C.; Null, A. P. In *Advances in Nucleic Acid and Protein Analyses, Manipulation, and Sequencing*, Limbach, P. A.; Owicki, J. C.; Raghavachari, R.; Tan, W., Eds.; Proc. SPIE 2000; San Jose, CA. Vol. 3926, pp 36–47. (e) Hannis, J. C.; Muddiman, D. C. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 954. (f) Hannis, J. C.; Muddiman, D. C., Fresenius J. *Anal. Chem.,* **2001,** *369*, 246–251. (g) Hannis, J. C.; Muddiman, D. C. *Rapid Commun. Mass Spectrom.*, **2001,** *15*, 348.
8. Demers, D. B.; Curry, E. T.; Egholm, M.; Sozer, A. C. *Nuc. Acids Res.* **1995,** *23*, 3050–3055.
9. (a) Greig, M.; Griffey, R. H. *Rapid Commun. Mass Spectrom.* **1995,** *9*, 97–102. (b) Muddiman, D. C.; Cheng, X. H.; Udseth, H. R.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1996,** *7*, 697–706.
10. (a) Hannis, J. C.; Muddiman, D. C. *J. Am. Soc. Mass Spectrom.* **2000,** *11*, 876. (b) Flora, J. W.; Muddiman, D. C. *Anal. Chem.* **2001,** *73*, 1247–1251.
11. Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D.-H.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 970–976
12. Ballistreri, A.; Garozzo, D.; Maravigna, P.; Montaudo, G.; Giuffrida, M. *J. Poly. Sci.: Part A: Polymer Chemistry* **1987,** *25*, 1049–1063 and references within.